基于知识与数据驱动的高速路车辆速度计算方法模型研究 *

彭大芹，罗裕枫，范兴荣

(1. “新一代信息网络与终端”重庆市协同创新中心，重庆 400065; 2. 重庆邮电大学 电子信息与网络工程研究院，重庆 400065)

摘　要：由于目前通过蜂窝基站定位的精度不够高和信令数据产生的不确定性，所以通过手机信令数据估计交通信息参数的精确度仍然是一个巨大的挑战。高速路车辆速度参数是评估道路服务质量的最常研究的参数之一，目前基于手机信令数据的交通参数估计基本步骤为数据收集及预处理、地图匹配、用户路线匹配确定和交通信息参数估计。在此基础上，根据物理运动学知识，针对手机信令数据产生的不确定性和容易参杂平行公路噪声用户数据的影响，提出一种融合距离加权的改进车辆速度计算方法，并在此方法之上进一步提出一种消除平行公路干扰数据的车辆速度计算方法，提高计算准确度。利用某移动公司提供的平台进行实例验证，并采用平均绝对误差 MAE、均方根误差 RMSE、平均绝对百分比误差 MAPE 作为性能指标，该方法模型性能均优于传统方法模型，并且平均绝对百分比误差比传统方法降低 10%左右，具有重要的应用价值。

关键词：智能交通系统；手机信令数据；车辆速度；平行道路；噪声数据

中图分类号：U　　doi: 10.3969/j.issn.1001-3695.2018.03.0155

Research on model of highway vehicle speed calculation based on knowledge and data driven

Peng Daqin, Luo Yufeng, Fan Xingrong

(1. *Chongqing Collaborative Innovation Center for Information Communication Technology*, *Chongqing* 400065, *China*; 2. *Electronic Information & Networking Research Institute*, *Chongqing University of Posts & Telecommunications*, *Chongqing* 400065, *China*)

Abstract: Due to the current high accuracy of positioning via cellular base stations and the uncertainty of signaling data, estimating the accuracy of traffic information parameters using cell phone signaling data remains a significant challenge. Highway speed parameters are one of the most commonly studied parameters for evaluating road service quality. At present, the basic steps of traffic parameter estimation based on cell phone signaling data are as follows: Data collection and preprocessing, map matching, user route matching determination and traffic information parameter estimation. On this basis, according to the knowledge of physical kinematics, aiming at the uncertainty of handset signaling data and the influence of parallel highway noise on user data, an improved vehicle speed calculation method based on weighted distance fusion is proposed and further proposed on this method A Method of Vehicle Speed Calculation to Eliminate Parallel Highway Interference Data and Improve Calculation Precision. Using a platform provided by a mobile company for instance verification, and using the average absolute error MAE, root mean square error RMSE, average absolute percentage error MAPE as a performance indicator, in this paper, the performance of the method model is better than the traditional method model, and the average absolute percentage error is about 10% lower than the traditional method, which has important application value.

Key words: intelligent transportation system; cell phone signaling data; Vehicle speed; Parallel road; Noise data

0　引言

近年来，经济的增长和技术进步的崛起导致了智能交通系统（ITS）对交通服务的需求越来越高，如何构建 ITS 的实时交通信息系统越来越重要。目前高速路的交通流数据主要通过感应线圈、地磁、视频、雷达探测器以及红外等检测器，这些方式需要在道路上安装检测器设备，需要耗费大量的人力物力。或者通过 GPS 进行车辆信息采集，但这种方式需要在运行车辆

收稿日期：2018-03-10；**修回日期：**2018-05-03　　**基金项目：**工信部新一代宽带无线移动通信网重大专项资助项目（2016ZX03002019-007）

作者简介：彭大芹（1969-），男，四川雅安人，正高级工程师，硕导，主要研究方向为智能终端、移动互联网、移动大数据等（pengdq@cqupt.edu.cn）；罗裕枫（1992-），男，广西梧州人，硕士研究生，主要研究方向为移动大数据；范兴荣（1986-），男，重庆人，博士研究生，主要研究方向为大数据算法、机器学习、模式识别与智能系统等.

上装载 GPS 相关设备，初期投资成本高，并且收集数据不完整，具有一定的局限性。随着蜂窝网络的全面密集覆盖和手机的普及，使用手机信令数据来获取交通信息参数在目前已经成为一种有效可行的方式，由于其不仅成本底和全面覆盖，还可以准确实时全天候获取交通参数信息，所以基于蜂窝手机信令数据估计交通参数和状况已经成为了一种替代传统方式的热门的技术方案。

近年来，有很多关于通过移动手机信令数据来获取高速公路交通流参数和路况信息方面的国内外的研究，主要的交通流参数包括车辆速度、车辆密度、行程时间和车流量等[错误!未找到引用源。]，其中车辆速度参数是高速公路交通信息中最重要参数之一，也是在评估道路服务质量的最常研究的参数之一。目前基于移动手机信令数据的高速公路车辆速度计算方法主要分为两种：第一种是通过两点的行程距离差与行程时间差的比值$V_k=\frac{S_2-S_1}{T_2-T_1}$来计算速度；第二种是通过路段的交通流量（veh/h）与交通密度（veh/km）的比值$V_k=\frac{Q_i}{K_i}$来计算速度。文献**错误!未找到引用源。**中基于对手机信令数据的分析验证了信令数据在交通拥堵识别方面的重要价值，并通过手机用户在行驶过程中在基站间切换产生的信令数据来计算行程距离差和行程时间来计算车辆速度，进而进行拥堵状态识别，证明了这种方法的可行性和实用性。当手机用户执行呼叫并从小区1移动到小区2时，产生并记录越区切换信号。此外，手机用户保持从小区2移动到小区3，还生成并记录另一越区切换信号。可以分析这两个越区切换信号以获得小区2的小区驻留时间和车辆行程距离。然后可以估计高速公路路段的车辆速度[错误!未找到引用源。]。文献**错误!未找到引用源。**提出基于位置切换的交通速度估算方法，根据用户轨迹与目标路段切换序列使用求最长公共子序列的方法进行地图匹配，利用切换点距离与行程时间的比值来估算交通流速度，并取得了不错的效果，但速度波动较大，存在干扰数据，致使计算精度不够高。文献**错误!未找到引用源。**通过手机用户位置切换和正常位置更新的数量估计交通密度，并根据用户呼叫到达和周期性更新的数量估计交通流量，进而估算车辆速度，证明了方法可行，但是由于利用手机信令数据估计交通密度和交通流量的准确度不够高和计算复杂，故存在较大误差。后来研究者发现利用手机信令数据估计交通速度会存在大量噪声数据的影响，如高速路周边的行人、周边公路的车辆用户等，并有研究者提出对信令数据进行分类，区分机动车载体和非机动车载体，同时滤除噪声数据的方法，通过此方法计算的车辆速度准确度得到了提高[错误!未找到引用源。]。文献**错误!未找到引用源。**提出了一种基于手机信令数据的高速路车速估计算法，引入了信令间隔区间阈值和时间间隔阈值，通过该阈值的限制，能够滤除一部分误差数据，能很大程度降低定位不准确导致的速度计算误差。

综合上述研究，由于手机信令数据产生的不确定性和容易参杂噪声数据的缺点，致使高速公路车辆速度计算准确度不够高。本文针对信令数据的产生的不确定性而在一个高速路段生成的距离长度不一的路程差的特点，提出利用两点的行程距离差与行程时间差的比值中引入与距离相关的加权因子函数来计算车辆速度的计算方法。还特别针对高速路周边平行公路干扰数据，进一步提出一种基于移动手机信令数据的平行公路干扰消除的车辆速度计算方法，很大程度提高了高速路车辆速度计算的准确度，对于基于手机信令手机获取交通信息的方法可以很好应用在现代智能交通系统（ITS），进一步提高服务质量具有重要的意义。

1　基本概念

1.1　产生手机信令数据的通信事件

当车载手机用户在高速公路上行驶的过程中会不确定性的在周边基站产生手机信令数据，手机信令数据的产生是由一些通信事件所驱动生成的，以下是产生手机信令数据的通信事件的一个简单的概括介绍，如图 1（b）所示。

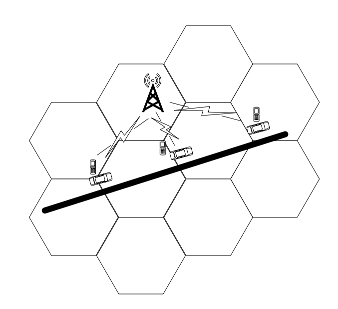


图 1　(a)车载用户在高速路上行驶过程

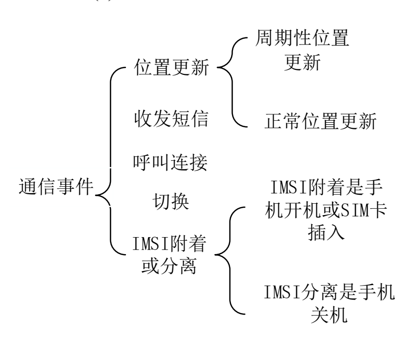


图 1　(b)通信事件简介图

由以上 5 类通信事件产生的手机信令数据经过移动运营商的手机原始信令数据采集系统平台并经过对数据进行过滤去噪等预处理后，再通过数据传输管道分发到各个处理模块。提取接收 G85 高速公路监测区域的基站的手机信令数据用于本文的实验研究，手机信令数据包含的主要字段与说明如表 1 所示。本文主要用到 1，2，3，4，5 的字段，其中通过 LAC 和 CELLID 可以匹配到基站的经度（Lng）和纬度（Lat）。

表 1　移动用户信令数据部分字段说明

序号	字段名称	字段说明
1	IMSI	国际移动用户识别码，唯一的手机识别
2	TimeStamp	信令数据发生的时间戳，YYYYMMDDHHMMSS，精确到秒

3	LAC	位置区编号
4	CELLID	小区位置区编号
5	EventID	通信事件类型：（1 周期性更新，2 正常位置更新，3 收发短信，4 呼叫切换，5 切换，6 开关机）
6	Flag	进出小区标识
7	Cause	切换原因
8	Result	事件结果

1.2　手机信令数据的噪声数据

高速公路路网复杂，临近周边存在平行公路、乡村道路和省道路等，并且由于基站的定位精度不够高，所以原始手机信令数据中会引入大量的噪声用户数据，例如高速路周边道路的行人、周边的居民、高速路临近平行公路和乡村道路的用户等噪声用户数据。这些干扰噪声数据会在车辆速度计算时造成严重的影响，导致车辆速度计算的不准确。目前针对高速路周边居民和行人的干扰数据已经存在解决的方法，但是没有特别针对高速公路临近平行公路和乡村道路等的车辆用户干扰数据的解决方法。因此本文提出一种平行公路干扰消除的方法，并在此基础上进行车辆速度计算。高速路周边的平行道路图如图 2 所示。

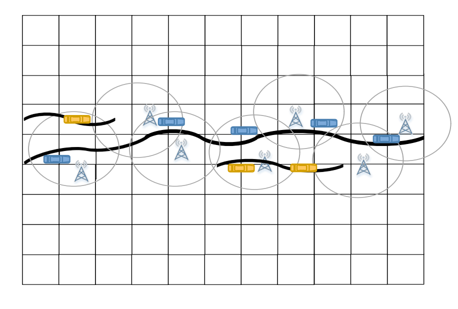


图 2　平行道路图

2　高速路车辆速度计算方法原理

目前，对于利用手机信令数据获取高速公路交通流参数信息已经存在了一定深入的研究，特别的，研究者对于基于手机信令数据的高速公路车辆速度计算方法的研究也做出了巨大的贡献[6~11]，为了从获取的手机信令数据中获取到高速路的车辆速度参数，可以包括以下基本步骤：信令数据预处理；路网匹配，进行高速路和基站序列的匹配；高速公路用户路径匹配，判断识别高速公路行驶车载用户；高速公路车辆速度计算。具体的流程框架图如图 3 所示。

2.1　数据清洗与预处理

由于基站定位技术的定位精度不够高、高速路周边环境的复杂性和手机信令数据产生的不确定性，由移动运营商的手机原始信令数据采集系统平台收集的信令数据会存在大量的“脏数据”，所以必须对收集的原始手机信令数据进行字段提取、滤除脏数据、过滤乒乓效应数据等数据预处理操作，有利于下一步的快速计算，提高匹配精度。其数据预处理的处理步骤如图 4 所示。

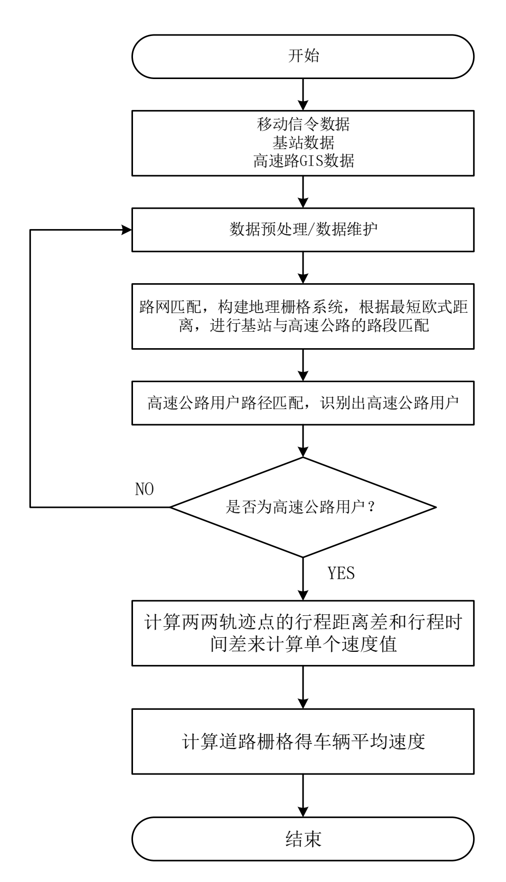


图 3　传统高速路车辆速度计算方法流程图

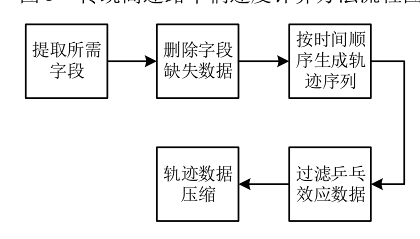


图 4　数据预处理

2.2　路网匹配

路网匹配也就是高速公路路段与移动基站覆盖的匹配，根据高速公路路段序列建立与基站的映射关系表是进行高速公路用户路径匹配和车辆速度计算的基础前提，本文考虑到道路段与基站的距离和基站的扇区覆盖方向等影响因素进行高速路路段与基站的最大可能匹配，其主要的步骤如下：

a)根据高速公路道路的路线，建立一个矩形监测区域，使得矩形区域完全包括高速公路，以矩形区域的左下角点为原点建立二维坐标系，并以长为 L 的正方形栅格对矩形区域进行栅格划分。记录高速路为道路栅格序列

$$\mathrm{Seg} = \{S_{x_1,y_1}(\mathrm{lng},\mathrm{lat}),\dots,S_{x_i,y_i}(\mathrm{lng},\mathrm{lat}),\dots,S_{x_n,y_n}(\mathrm{lng},\mathrm{lat})\}$$

记录监测区域的基站序列为

$$\mathrm{sta} = \{M_{x_1,y_1}(\mathrm{lng},\mathrm{lat}),\dots,M_{x_i,y_i}(\mathrm{lng},\mathrm{lat}),\dots,M_{x_n,y_n}(\mathrm{lng},\mathrm{lat})\}$$

b)为每个道路栅格最大可能匹配对应的基站。通过经纬度计算道路栅格序列里每个栅格与基站序列的每个基站的欧式距离，并且确定基站扇区是否覆盖到道路栅格，选取距离最短并且基站扇区覆盖道路栅格的基站为该道路栅格的映射基站。

c)以二维坐标序列号升序方向$\{(x_1,y_1),..(x_i,y_i),..(x_n,y_n)\}$为高速路的正方向，降序方向$\{(x_n,y_n),..(x_i,y_i),..(x_1,y_1)\}$为高

速路反方向，建立高速路段和基站的匹配库，如表 2 所示。

表 2 高速路与基站匹配说明表

序号	字段名	字段说明
1	Road_ID	高速路道路栅格序号(x_i,y_i)
2	Road_Lng	道路栅格经度
3	Road_Lat	道路栅格纬度
4	Station_ID	基站栅格序号(x_i,y_i)
5	Station_Lng	基站经度
6	Station_Lat	基站纬度
7	Flag	基站是否覆盖道路栅格，是：1，否：0

2.3 高速公路用户路径匹配

高速公路用户路径匹配是为了将经过数据预处理后的用户手机信令数据轨迹序列与高速路栅格序列进行匹配，初步识别出在高速路上行驶的车载用户。其主要的步骤如下：

a)获取经过数据预处理得到的监测区域的用户手机信令数据轨迹序列集，提取每个用户的信令轨迹序列$Trace_i=\{M_{x_jy_j}(lng,lat),\ldots,M_{x_ky_k}(lng,lat)\}$共有 N 个轨迹点，将用户的信令轨迹序列与高速公路基站匹配库进行对比，统计用户的信令轨迹序列在高速公路基站匹配库中存在的N_1个轨迹点。

b)计算用户信令轨迹与高速路道路栅格序列的相似度 λ $=\frac{N_1}{N}$，当相似度 λ 超过一定阈值并且具有一定方向性则判定为高速公路用户，用户的方向性通过二维坐标序列号(x_i,y_i)的递增和递减来确定。

2.4 高速公路车辆速度计算

文献[11]传统高速公路道路栅格车辆用户速度的计算主要是通过两轨迹点的行程路程差和行程时间差的比值来计算。

$$V_i=\frac{dist\{M_{x_iy_i}(lng,lat),M_{x_jy_j}(lng,lat)\}}{T_j-T_i} \quad (1)$$

其中:$M_{x_iy_i}(lng,lat)$是基站位置点，dist 是计算两个基站位置点之间的欧式距离，T_j-T_i是两点的行程时间差。

若计算道路栅格 K 的平均速度，假设有 n 个两两轨迹点经过并包括道路栅格 K，则道路栅格 K 的平均车辆速度为V_k。

$$V_k=\sum_{i=1}^{n}{}^{V_i}/_{n} \quad (2)$$

1) 本文方法模型 1

传统方法单纯考虑了行程路程差和行程时间差的比值来计算车辆速度，但由于手机信令数据产生的不确定性，进而会造成包含一个道路栅格会存在行程距离长度不一样的多个两两轨迹点。这种不确定性会对高速路车辆速度计算有一定的影响，因此文本在传统方法的基础上，根据物理运动学的思想知识，由统计意义的角度上可知包含一个道路栅格的两轨迹点计算单元的路程距离越短，则对该道路栅格平均速度的贡献值越大的特点，因此本文根据此特点提出一种融合路程加权的改进车辆速度计算方法，其改进方法的具体流程框图如图 5 所示。由流程图可知改进部分方法步骤如下：

a) 计算包含某道路栅格的两两轨迹点的行程距离和行程时间，用于计算对该道路栅格的车辆平均速度的速度贡献值的其中之一。

b)选择一个与行程距离呈反相关的加权因子函数φ_{ij}^{l}计算加权值，用于对速度贡献值进行加权。

c)使用距离权值函数计算加权值然后与每个速度贡献值的乘积后的值进行求和运算得到最终的道路栅格车辆平均速度值。计算公式如下：

$$V_k=\sum_{l=1}^{N_k}\varphi_{ij}^{l}\cdot\frac{dist\{M_{x_iy_i}(lng,lat),M_{x_jy_j}(lng,lat)\}}{T_j-T_i} \quad (3)$$

其中l表示为第 k 个路段的第l个用户信令，N_k表示第 k 个路段的两两轨迹点数。

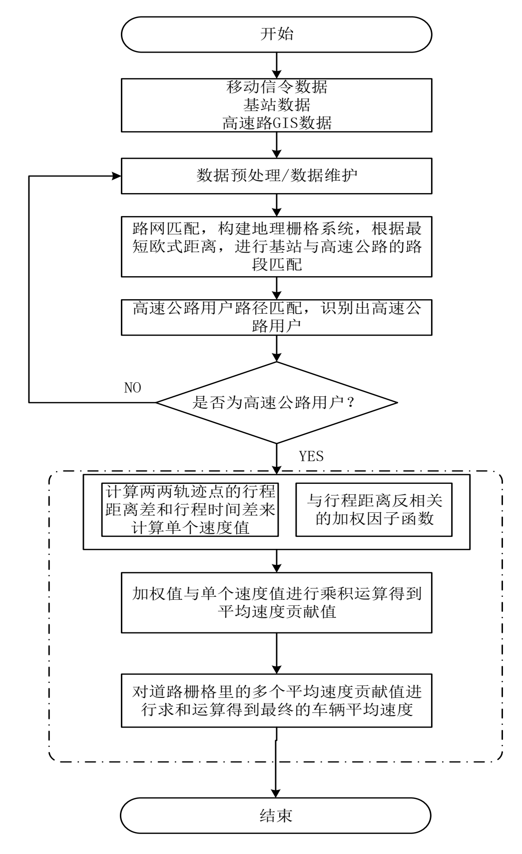


图 5 本文方法模型 1 方法流程

2) 本文方法模型 2

高速公路车辆速度计算准确度除了受手机信令数据产生不确定性特点的影响外，还很大程度受到了噪声用户数据的干扰，特别是高速路临近的平行公路、乡村道路和省道路的行车用户的影响。由于高速路的平行道路一般存在于高速路的路段周边，因此可以根据其前面不存在平行道路的路段的历史用户数据进行剔除干扰用户。因此针对平行道路干扰用户，在本文方法模型一改进的车辆速度计算方法的基础上进一步对平行公路干扰消除，然后重新计算含平行道路的高速路车辆速度。主要依据是存在平行公路路段的高速路与之前高速路段的用户数量差别和高速路与平行公路之间的速度差别的思想进行干扰剔除，其具体的流程框图如图 6 所示。其方法主要的原理如下：

d)获取当前时刻的存在平行公路的高速路道路栅格的用户数N_k个用户，并同时获取其历史时刻的后方不存在平行公路的高速路道路栅格的用户数N_t个用户，用户数N_k与用户数N_t求交

集运算$N_k^1=N_k\cap N_t$，那么$N_k^2=N_k-N_k^1$，则N_k^1部分用户是确定为高速公路上的用户，N_k^2部分用户是不确定的用户；

e)根据确定为高速公路上的部分用户N_k^1来计算存在平行公路的道路栅格的平均速度值V_k^1，平均速度值V_k^1运用本文方法模型一的车辆平均速度计算方法来计算。然后根据平均速度值V_k^1来确定该道路栅格属于那种交通状态，交通状态分为畅通状态（$V_k^1\geq 80km/h$），轻度拥堵状态（$40\ \mathrm{km/h}\leq V_k^1<80\ \mathrm{km/h}$），拥堵状态（$V_k^1<40\ \mathrm{km/h}$）。

f)根据行程路程差和行程时间差的比值的速度计算公式计算N_k^2中每个用户的速度值V_k^2，然后与速度值V_k^1进行对比，如果某用户的速度值V_k^2属于由V_k^1确定的道路状态的速度范围，则判定为高速公路上的用户，否则判定为平行公路干扰用户，并进行剔除处理。并且得到从不确定为高速公路用户的部分用户N_k^2中进一步确定为高速路上的用户N_k^3。

g)对N_k^3部分用户再利用本文方法模型 1 的车辆速度计算方法计算平均速度值V_k^3。

h)根据经过平行公路噪声数据用户消除后的用户数据重新进行道路栅格的车辆平均速度的计算。计算公式如下：

$$V_k=\frac{N_k^1}{N_k^1+N_k^3}\times V_k^1+\frac{N_k^3}{N_k^1+N_k^3}\times V_k^3 \tag{4}$$

开始
移动信令数据
基站数据
高速路GIS数据
数据预处理/数据维护
路网匹配，构建地理栅格系统，根据最短欧式距离，进行基站与高速公路的路段匹配
高速公路用户路径匹配，识别出高速公路用户
是否为高速公路用户？
NO
YES
计算两两轨迹点的行程距离差和行程时间差来计算单个速度值
与行程距离反相关的加权权值函数
加权值与单个速度值进行乘积运算得到平均速度贡献值
对道路栅格里的多个平均速度贡献值进行求和运算得到路段栅格最终的车辆平均速度
当前平行公路的高速路道路栅格用户与历史时刻的前面不存在平行公路高速路道路栅格用户进行交集运算
确定高速公路部分用户
不确定部分用户
将确定部分用户，按照本文方法模型一计算速度V，确定交通状态
对不确定部分用户，计算每个用户的速度值Vi
Vi是否属于由V确定的交通状态的速度范围
NO
剔除用户
YES
按照用户数加权的速度计算方法重新计算平行路段的高速路道路栅格车辆平均速度
结束

图 6　本文方法模型 2 方法流程

3　实验

3.1　实验条件介绍

本文选择重庆市 G85 成渝高速西环立交至永川收费站路段作为模型的实验试点路段，路段全长 56 公里，使用矩形框区域对路段进行对路段的框定，如图 6 所示。



图 6　G85 高速路测试路段

手机信令数据采集系统实时全天候采集试点高速路段周边的基站的移动信令数据并通过 ActiveMQ 消息机制实时分发到模型进行计算处理。其采集的原始移动信令数据的示例如图 7 所示：



图 7　移动信令数据原始数据示例图

由图 7 可知，原始数据中主要包括了用户 ID、产生时间、基站位置、进去小区标准和切换原因等字段，特别的，jumpin 表示手机用户进入小区覆盖区，jumpout 则表示用户离开小区覆盖区，move 表示在小区覆盖区中移动。

以移动信令数据为数据输入源，模型结果的验证采用安装在高速路旁的固定检测器数据进行与模型的输出结果数据验证比对，为了验证消除平行公路噪声数据方法模型的有效性，选择了具有平行公路的高速路路段的两个固定检测器，固定检测器可检测每 5 分钟内的瞬时平均速度。模型的性能指标采用平均绝对误差 MAE，均方根误差 RMSE，平均绝对百分比误差 MAPE，计算公式如下：

$$\mathrm{MAE}=\frac{1}{N}\sum_{i=1}^{N}\left|x_i-x_i^{'}\right| \tag{5}$$

$$\mathrm{RMSE}=\sqrt{\frac{\sum_{i=1}^{N}(x_i-x_i^{'})^2}{N}} \tag{6}$$

$$\mathrm{MAPE}=\frac{1}{N}\sum_{i=1}^{N}\left|\frac{x_i-x_i^{'}}{x_i}\right| \tag{7}$$

其中:N 为样本量数；x_i为检测器的结果；x_i'为模型计算的结果。

3.2　结果验证分析

测试路段监测区域的移动信令数据流由某移动运营商提供，数据流经过 ActiveMQ 消息机制进行实时分发，每接收五分钟的信令数据后分发给数据处理模型经过 2.1、2.2、2.3 节的步骤处理后的数据再分别经过如下三种高速路车辆平均速度计算模型进行处理得到路段的平均车辆速度：a)文献[11]模型：传统车

辆平均速度计算模型；b)本文方法模型 1：融合路程加权车辆平均速度计算模型；c)本文方法模型 2：融合距离加权+平行公路干扰数据消除车辆平均速度计算模型。

在 2.4 节的本文方法模型 1 中，为了取得最好的计算结果，需要选择一个最合适的路程反相关的加权函数，常用的加权函数有路程倒数函数$\frac{1}{S}$、路程平方倒数函数$\frac{1}{S^2}$、高斯加权函数$e^{-(S/_{\delta})^2}$，其中 S 是指车辆行程距离。将 2017 年 5 月 17 日 6：00 至 22：00 的数据分别输入到使用以上的三种加权函数的车辆平均速度计算方法模型和传统的车辆平均速度计算模型进行处理，取两个固定速度检测器所在的路段模型的计算结果与固定速度检测器数据进行对比分析，得到误差性能对比表如表 3 所示。

表 3 本文方法模型 1 性能误差对比表

模型名称	速度检测器	MAE	RMSE	MAPE
文献[11]方法模型	1	16.312	19.247	0.187
	2	17.584	20.329	0.193
本文方法模型 1（$\frac{1}{S}$加权）	1	8.945	11.643	0.122
	2	9.912	13.642	0.135
本文方法模型 1（$\frac{1}{S^2}$加权）	1	8.547	11.476	0.113
	2	8.957	12.243	0.121
本文方法模型 1（$e^{-(S/_{\delta})^2}$加权）	1	7.983	10.671	0.105
	2	8.256	10.968	0.110

从表 3 可以看出，本文方法模型 1 的计算结果都优于文献[11]方法模型结果，验证了本文方法模型 1 提出的融合路程加权的改进车辆速度计算方法的有效性，并且使用 $e^{-(S/_{\delta})^2}$的加权模型的计算结果最优，因此选择$e^{-(S/_{\delta})^2}$作为加权函数。

进一步考虑平行公路干扰数据的影响，在融合路程加权的车辆平均速度计算模型的基础上加入平行公路干扰数据消除方法模型得到融合路程加权+平行公路干扰消除的车辆平均速度计算模型，本文命名为方法模型 2，将该模型结果和本文方法模型 1 的结果与两个固定测速仪数据进行对比分析，得到误差性能对比如表 4 所示。

表 4 本文方法模型二性能误差对比表

模型名称	速度检测器	MAE	RMSE	MAPE
本文方法模型 1	1	7.983	10.671	0.105
	2	8.256	10.968	0.110
本文方法模型 2	1	6.838	8.404	0.086
	2	7.015	8.517	0.091

从表 4 可知，在加入了平行公路干扰消除后，模型的车辆速度计算结果的 MAE、RMES 和 MAPE 都得到了进一步的提升，该方法能有效的消除平行公路的噪声用户数据，避免了高速路临近的平行道路的噪声用户对高速路车辆速度计算的影响。

在选定了加权因子函数为高斯加权函数后，给出 2017 年 5 月 17 日 6：00 至 22：00 的本文方法模型 1、本文方法模型 2 与测速仪 1 的车辆平均速度结果对比图和误差分布图，分别如下图 8、9 所示，本文方法模型 1、本文方法模型 2 与测速仪 2 的车辆平均速度结果对比图和误差分布图，分别如下图 10、11 所示，从图 8、10 中可以看出经过平行道路噪声数据消除的本文模型 2 的结果跳动幅度变小，更加接近测速仪的结果。从误差分布图 9 可以看出本文方法模型 1 相对误差 10%以内的占比为 74%，相对误差 10%~20%以内占比为 23%；本文方法模型 2 相对误差 10%以内占比为 84%，相对误差 10%~15%以内占比为 16%；从误差分布图 11 可以看出本文方法模型 1 相对误差 10%以内的占比为 68%，相对误差 10%~20%以内占比为 26%；本文方法模型 2 相对误差 10%以内占比为 75%，相对误差 10%~15%以内占比为 25%，可以明显看出，本文方法模型 2 相对本文方法模型 1 消除了部分 20%以上的误差占比，结果更为准确，进一步证明了本文方法模型 2 对于平行公路噪声数据消除的有效性。综合来说，本文方法模型 2 的计算结果平均绝对误差为 9%左右，比本文方法模型 1 的平均绝对误差降低了 2%左右，比文献[11]方法模型的平均绝对误差降低了 10%左右。

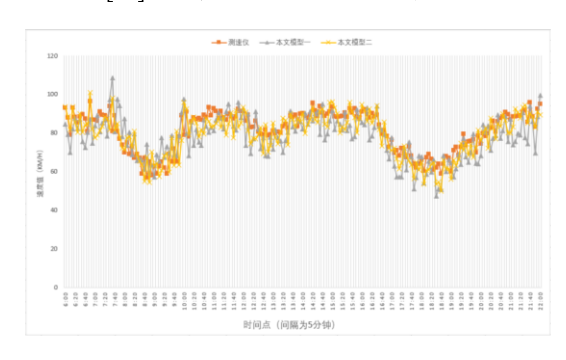
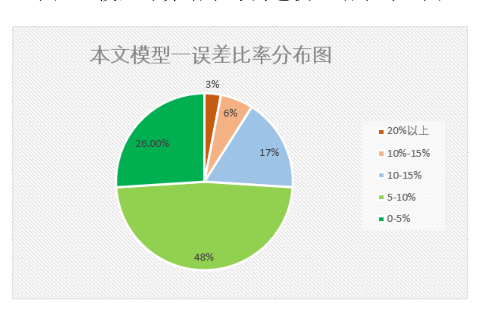
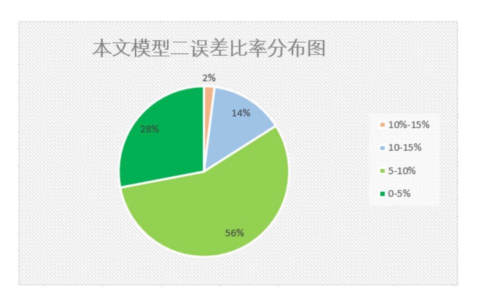


图 8 模型计算结果与测速仪 1 结果对比图



(a)本文方法模型 1



(b)本文方法模型 2

图 9 模型误差分布图

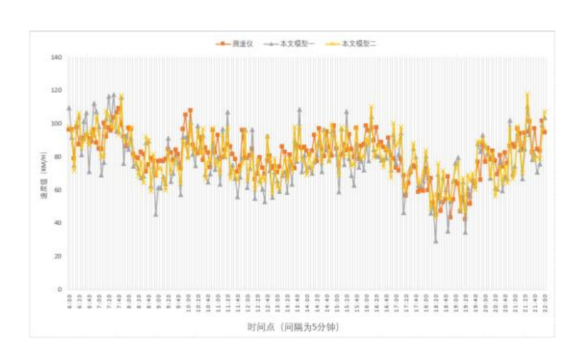
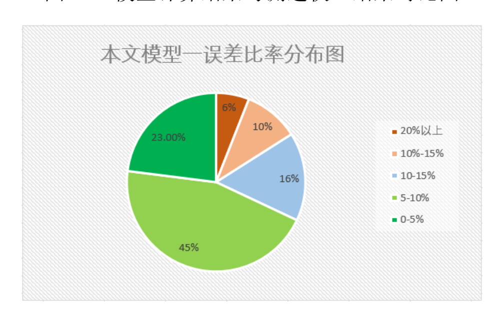
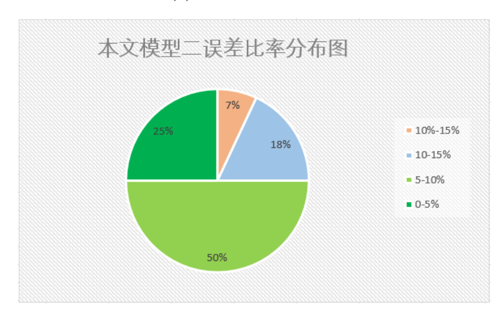


图 10 模型计算结果与测速仪 2 结果对比图



(a)本文方法模型 1



(b)本文方法模型 2

图 11 模型误差分布图

4 结束语

利用手机信令数据获取高速路道路交通信息对于智能交通系统的发展具有重要的意义。本文简述了移动手机信令数据的特点和基于移动手机信令数据获取高速公路车辆速度基本方法原理，根据物理运动学知识，针对手机信令数据产生不确定性和容易参杂噪声数据特别是平行道路噪声数据的特点，分别提出了一种融合路程加权的车辆速度计算方法模型，并在此方法之上进一步提出一种消除平行公路干扰数据的车辆速度计算方法。以 MAE、RMSE 和 MAPE 为性能指标与文献[11]传统方法模型进行相互的对比分析验证，选定高斯加权函数作为加权因子函数，并证明了本文方法模型的性能指标均优于传统方法模型。本文方法对于促进智能交通系统的发展具有重要的应用价值，并且如果需要大规模的实现，可考虑使用相关的大数据平台进行实现[12]。

参考文献：

[1] Caceres N, Wideberg J P, Benitez F G. Review of traffic data estimations extracted from cellular networks [J]. IET Intelligent Transport Systems, 2008, 2 (3): 179-192.

[2] 沈激，姜新农，周一. 基于手机信令的高速公路拥堵信息自动识别 [J]. 中国交通信息化, 2015 (S1): 119-122. (Shen Lian, Jiang Xinnong, Zhou Yi. Automatic identification of freeway congestion information based on mobile signaling [J]. Transportation Information Industry, 2015 (S1): 119-122.)

[3] Wu Chei, Chen Chihua, Lin Bonyeh, *et al.* Traffic information estimation methods from handover events [J]. Journal of Testing & Evaluation, 2016, 44 (1S): 656-664.

[4] Chang Mingfeng, Chen Chihua, Lin Yibing, *et al.* The frequency of CFVD speed report for highway traffic [J]. Wireless Communications & Mobile Computing, 2015, 15 (5): 879-888.

[5] Demissie M G, Correia G H D A, Bento C. Intelligent road traffic status detection system through cellular networks handover information: An exploratory study [J]. Transportation Research, Part C, 2013, 32 (4): 76-88.

[6] Shen Yun, Dong Honghui, Jia Limin, *et al.* A method of traffic travel status segmentation based on position trajectories [C]// Proc of the 18th IEEE International Conference on Intelligent Transportation Systems. 2015: 2877-2882.

[7] Chen Chihua, Chang Hsuchia, Su Chunyun, *et al.* Traffic speed estimation based on normal location updates and call arrivals from cellular networks [J]. Simulation Modelling Practice & Theory, 2013, 35 (6): 26-33.

[8] Chang Hsuchia, Chen Chihua, Lin Bonyeh, *et al.* Traffic information estimation using periodic location update events [J]. International Journal of Innovative Computing Information & Control Ijicic, 2013, 9 (5): 2031-2041.

[9] Lai Weikuang, Kuo Tinguan, Chen Chihua. Vehicle speed estimation and forecasting methods based on cellular floating vehicle data [J]. Applied Sciences, 2016, 6 (2): 47.

[10] Zhu Tongyu, Song Zhixin, Wu Dongdong, *et al.* A novel freeway traffic speed estimation model with massive cellular signaling data [J]. International Journal of Web Services Research, 2016, 13 (1): 69-87.

[11] 刚红润，沈志纲，刘承华，等. 一种基于手机信令的高速路车速估计算法 [J]. 电子世界, 2017 (17): 46-46. (Gang Hongrun, Shen ZhiGang, Liu Chenghua, *et al.* An algorithm of highway speed estimation based on cellular signaling [J]. Electronics World, 2017 (17): 46-46.)

[12] 杨杰，朱邦培，吴宏伟. 基于 Storm 的高速公路实时交通指数评估方法的研究与实现 [J]. 计算机应用研究, 2017 (9): 2707-2713. (Yang Jie, Zhu Bangpei, Wu Hongwei. Research and realization of evaluation method for freeway real-time traffic index based on storm [J]. Application Research of Computers, 2017 (9): 2707-2713.)